The Bell System Technical Journal

Vol. XXVIII *July, 1949* *No. 3*

Editorial Note regarding Semiconductors

ALL but one of the papers that comprise this issue discuss practical applications of semiconductors and touch upon their properties as employed in rectifying devices, detectors, and in a new amplifying unit—the so-called transistor. These semiconductor papers all relate to one another and present, as a whole, a current but well developed account of the behavior and uses of these very promising additions to today's vast array of electrical applicances.

Because semiconductors are relative newcomers, few engineers have as yet had occasion to become familiar with their characteristics and the reasons for their somewhat unexpected performance. Accordingly, it seems appropriate to preface the present group of papers with a brief introductory note devoted to the nature of the physical phenomena encountered.

The semiconductors of interest in the communications art are electronic rather than ionic conductors, and include copper oxide, various other oxides, selenium, germanium and silicon. Being electronic conductors, the constituent atoms remain in fixed positions. They may lose or gain electrons during the conduction process but the structure of the conductor as a whole and its chemical composition are not affected.

Basic to the theory of these semiconductors is the idea that electrons can carry current in two distinguishable and distinctly different ways: one being called "excess conduction," "conduction by excess electrons," or simply "conduction by electrons;" and the other being called "deficit conduction" or "conduction by holes." The possibility that these two processes may be simultaneously and separably active in a semiconductor affords a basis for explaining transistor action.

We shall confine our attention to the behavior of electrons within the silicon and germanium type of crystal lattice, and especially as it is modified by minute amounts of suitably chosen impurities.[1]

---

[1] There has been very marked development in the understanding of semiconductors since 1940. This understanding is an outgrowth of the research and development program on crystal rectifiers undertaken in connection with the radar program during the war and continued in several laboratories thereafter. Some of the wartime work was carried out in the Radiation Laboratory of M.I.T., which operated under the supervision of the National Defense Research Committee. The Radiation Laboratories Series volume "Crystal Rectifiers" by H. C. Torrey and C. A. Whitmer reports this program and mentions in particular as chief contributors to crystal research and development in England: the General Electric

Silicon and germanium form what are called "covalent crystals," the atoms being held together by "electron-pair bonds" formed by the valence electrons. The covalent bond in the hydrogen molecule is the simplest electron-pair bond. Figure 1 represents two hydrogen atoms and a hydrogen molecule.[2] Each atom consists of a proton and one electron. The proton weighs approximately 2,000 times as much as the electron and is a relatively immobile particle about which the electron moves in its orbit or quantum mechanical wave function. In an isolated atom, this wave function has spherical symmetry and the electronic charge is distributed on the average as a diffuse sphere centered about the proton. When the two atoms are brought close together, interaction between the wave functions of the two electrons takes place and the electronic cloud of each becomes modified, as suggested in the figure. The result is to produce an extra accumulation of charge between the two protons which acts to bind them together. According to quantum mechanical laws associated with the "Pauli exclusion principle," the bond is especially stable when it contains precisely two electrons. It is weakened considerably by removal of one electron and is not greatly strengthened by the addition of a third electron. This special stability of the electron-pair bond or covalent bond is a fundamental fact of chemistry which is now quite well understood on the basis of wave mechanics.

The elements carbon, silicon and germanium have the common feature of being tetravalent. Although they possess respectively 6, 14 and 32 electrons all together, in each case only four of these are able to enter into chemical reactions. The remaining electrons are closely bound to the nucleus producing a stable "ionic core" having a net charge of +4 units. This core can be regarded as completely inactive so far as electronic processes in chemical reactions and in semiconductors are concerned.

Each of these atoms tends to form covalent or electron-pair bonds with four other atoms. This tendency is completely satisfied in the diamond lattice which is the crystalline form of all three elements. The lattice, shown in Fig. 2, is a cubic arrangement and may be regarded as made up of four interpenetrating simple cubic lattices like the one formed by the atoms on the four corners of the cube shown. In this structure each typical atom is surrounded by four neighbors regularly placed about it, with which it forms four

---

Company, British Thompson-Houston Ltd., Telecommunications Research Establishment and Oxford University; and in the United States: the Bell Telephone Laboratories, Westinghouse Research Laboratory, General Electric Company, Sylvania Electric Products, Inc., and the E.I. duPont deNemours and Company. It is also pointed out that the crystal groups at the University of Pennsylvania and Purdue University, who operated under N.D.R.C. contracts, were responsible for much fundamental work.

[2] The figures in this introduction and the text associated with them, like the following paper on "Hole Injection in Germanium", follow closely the presentation in a book entitled "Holes and Electrons, an Introduction to the Physics of Transistors" now under preparation by W. Shockley.

covalent bonds. These neighbors are arranged on the corners of a regular tetrahedron in conformity with the known chemical behavior of the tetrahedral carbon atom.[3] For purposes of discussion of conductivity in these crystals, we shall represent the three-dimensional array in two dimensions as is shown in Fig. 3, indicating that each carbon atom forms an electron-pair bond with four neighbors.

On the basis of this valence bond structure we can intuitively see why diamond should be an insulator. Although it contains a large number of electrons, as does a metal, the covalent bond is a quite different structure from the metallic bond. In an ideally perfect diamond crystal, each valence bond would contain its two electrons; therefore, every electron would be tightly bound and thus unable to enter into the conduction process.

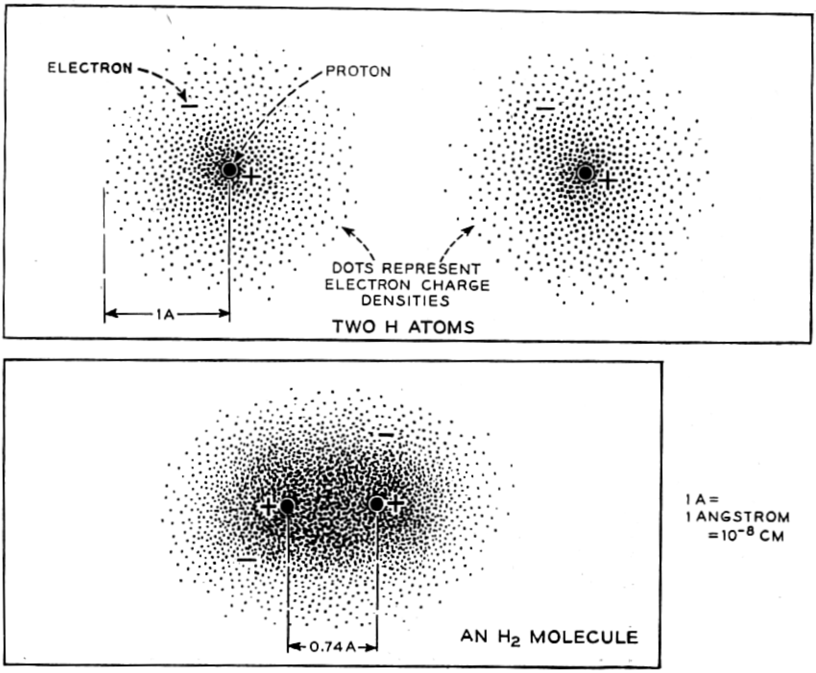


FIG. 1.

Conductivity can be produced in diamond, however, in a number of ways, all of which involve destroying the perfection of the valence bond structure.

---

[3] Long before the arrangement of atoms in the diamond crystal was established by X-rays, the organic chemists had concluded that carbon formed four bonds at the tetrahedral angles—a truly remarkable result of inductive reasoning based on observations of the optical properties of solutions of organic compounds.

Thus, if high-energy particles or quanta of radiation fall upon the crystal, they can break the bonds. Conductivity in diamond induced by bombardment in this way has recently received considerable prominence in connection with "crystal counters" which have been used to detect nuclear particles and in experiments on conductivity induced by electron bombardment. An electron ejected from a bond constitutes a localized negative charge in the crystal and, since initially the bond structure was electrically neutral, the electron as it departs from its point of liberation leaves behind an equal, localized positive charge. Such a migratory electron, because it represents an excess over and above that required to complete the bond structure in its neighborhood, is called an "excess electron." Since it cannot enter any of the completed bonds in the lattice, it moves about in the crystal in a random manner under the influence of thermal agitation. If an electric field is applied, it tends to drift in the direction of the applied force and to carry a current. This illustrates conduction by excess electrons (referred to simply as conduction by electrons) and, as we shall see, is to be distinguished from the other process whereby an electron deficit enables conduction to be effected by "holes."

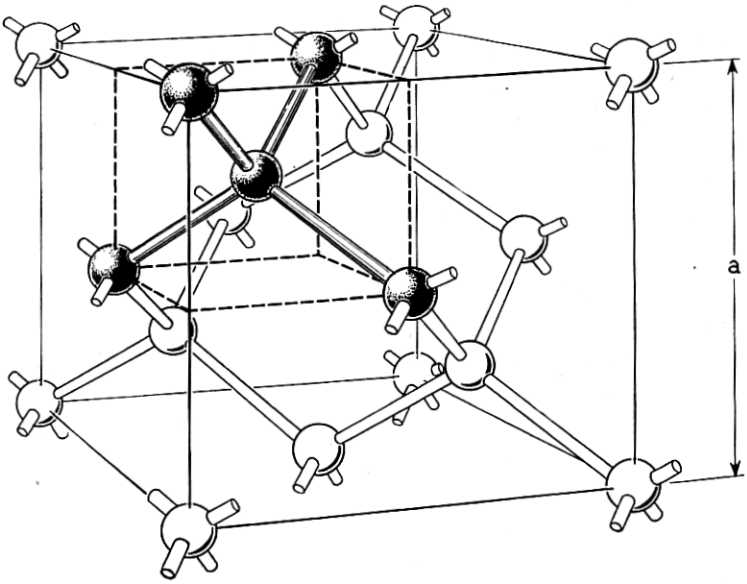
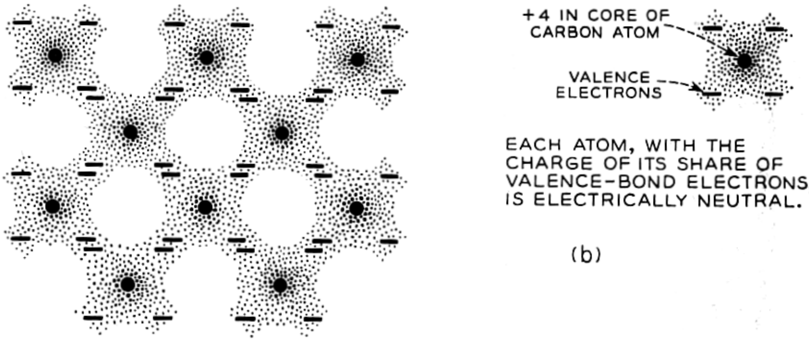


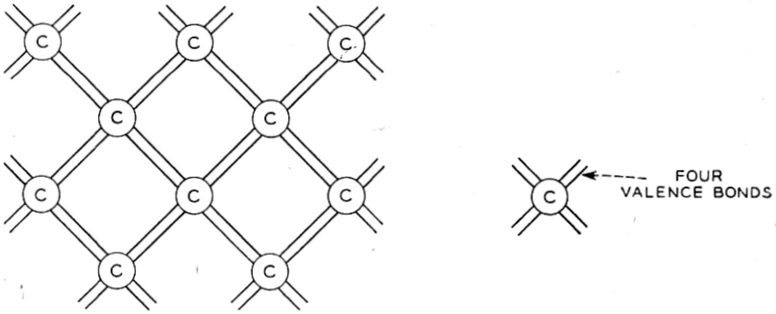
FIG. 2.

Such a hole, constituting a net, localized, positive charge in the crystal, moves from place to place by a reciprocal motion of electrons in the valence bonds; and, under the influence of an electric field, its random motion acquires a systematic drift. Therefore it also can contribute to the current; in other words, current can flow as well by virtue of a deficit of electrons as by an excess of them.

In an illuminated and bombarded diamond crystal the electrons and holes, produced in pairs by the excitation, will of course drift in opposite directions under the influence of a field; the electron, being negative, drifts in the opposite direction from the applied field, but its current is in the direction of the field. In the case of the hole, the reciprocal electron motions are once more opposite to the direction of the field (on the average). As a consequence, the net result is that the hole drifts in a direction to increase the current represented by the electrons. If the source of bombardment or illumination is removed, the conductivity dies away and the crystal will return to its normal state. This can occur by the recombination of holes and electrons. Whenever an electron drops into a hole, both the hole and the electron dis-



(a) ELECTRON PAIR BONDS



(C) PLANE DIAGRAM OF DIAMOND LATTICE WITH BONDS REPRESENTED BY LINES

FIG. 3.

appear and the bond structure becomes complete, the excess energy being given up to the atoms in the form of thermal vibrations.[4]

If the temperature is sufficiently elevated, spontaneous breaking of some fraction of the covalent bonds by agitation will occur producing electrons and holes in equal numbers. In a diamond this effect would occur at such high temperatures that it would not be observed. However, it plays a major role in silicon and germanium at temperatures well within the range of investigation in the laboratory.

On the basis of quantum mechanical theory, it is found that a very high degree of symmetry exists between the behavior of electrons and the behavior of holes. One may think of the hole as moving through the crystal as a positively charged particle with much the same attributes as a free electron except for the sign of its charge.

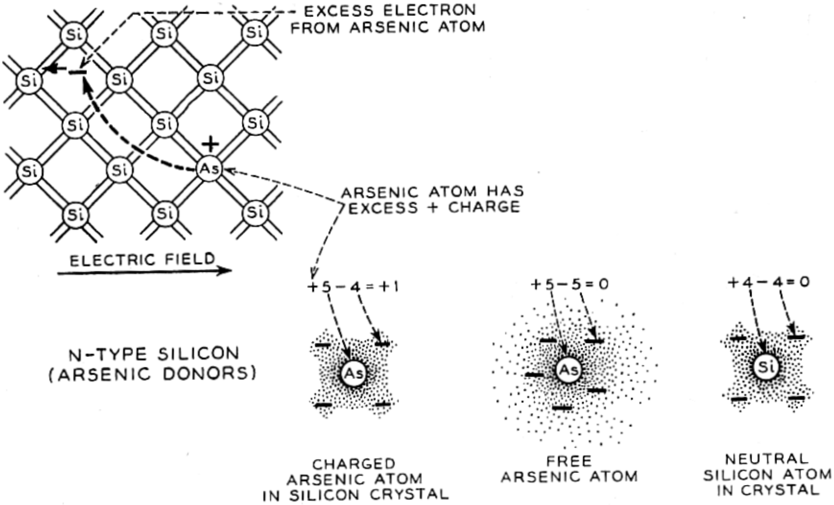


FIG. 4.

IMPURITY SEMICONDUCTORS: DONORS AND ACCEPTORS

If the only cases of conductivity open to investigation were like those discussed above, for which electrons and holes are present in equal numbers, the problem of interpreting the data would be very difficult. Fortunately, in the semiconductors silicon and germanium, there are cases in which conductivity is due to excess electrons only or to holes only.[5]

---

[4] The process of recombination may actually be much more complicated and may involve intermediate stages in which the hole or the electron is trapped.

[5] The behavior of silicon with impurities of the sorts discussed here was investigated by Scaff, Theurer, and Schumacher. Their work was stimulated by the development of silicon detectors for microwave use by R. S. Ohl, also of Bell Telephone Laboratories, in the prewar years.

If the conductivity of the sample is due to excess electrons it is called *n-type*, since the current carriers act like negative charges; if due to holes, it is called *p-type*, since the carriers act like positive charges.

Either type of conduction can be produced at will by admixture of a suitable "impurity," a donor such as arsenic yielding an excess of free electrons, while an acceptor like boron causes an electron deficit or a surplus of positive holes. The reason why arsenic and boron serve in these opposite capacities comes readily to hand.

The arsenic atom has five valence electrons surrounding a core having a net charge of +5 units and, when introduced (e.g. in silicon) as a low-fraction impurity, it is believed that each arsenic atom displaces one of the silicon atoms from its regular site and forms four covalent bonds with the nighboring silicon atoms, thus using four of its five valence electrons (see Fig. 4). The extra electron cannot fit into these four bonds and is free to move about the crystal. This excess electron therefore constitutes a mobile localized negative charge. The arsenic atom, on the other hand, is an *immobile* localized positive charge, since its core, with a charge of +5 units, is not neutralized by its share (−4) of the charge in the valence bonds. Its net charge, therefore, just balances that of the excess electron it sets free in the crystal. Thus arsenic impurity atoms add excess electrons but do not disturb the over-all electrical neutrality of the crystal. The negative electrons, however, are somewhat attracted by the positive arsenic atoms and at low temperatures become weakly bound to them. At room temperature, thermal agitation shakes them off so that they become free.

To produce a p-type semiconductor we choose an added impurity, such as boron, having three valence electrons and therefore not enough to complete the valence bond structure surrounding it. The hole in one of the bonds to the boron atom can be filled by an electron from an adjacent bond, and when this occurs the hole migrates away to the bond which just gave up one of its electrons. The boron atom thus becomes an *immobile* localized negative charge. Because of the symmetry between the behavior of holes and electrons, we can describe the situation by saying that the negative boron atom attracts the positively charged hole but that thermal agitation shakes the latter off at room temperature so that it is free to wander about and contribute to the conductivity.

Because of their valencies, phosphorous and antimony, as well as arsenic are in the donor class while aluminum, gallium and indium are additional examples of the acceptor class.

It is beyond the scope of this prefatory note to describe how, by measurements of conductivity and the Hall effect as influenced by the amount of added donor or acceptor, it has been possible to determine the concentration of electrons and holes, as well as to fix the energies needed to remove an electron from a donor, a hole from an acceptor, and to break a covalent bond between lattice atoms. In samples of germanium of such purity that the amount of added donor or acceptor was too small to determine by conventional chemical methods, the conductivity was still controlled by the processes outlined above. And it is interesting to note that a portion of this investigation was carried out with the aid of radioactive antimony alloyed with the germanium, the radioactive property making possible an accurate count of antimony atoms, though present only in extremely attenuated amounts.

The semiconductor papers in this issue of the Journal will explain how these simple facts of electron exchange give rise to rectifying and amplifying properties.

SEMICONDUCTOR RECTIFIERS AND AMPLIFIERS

A contact between a metal and semiconductor may act as a rectifier, the contact resistance being high for one direction of current flow and low for the opposite. Rectification results from the presence in the semiconductor adjacent to the interface of a potential barrier or hill which the current carriers, electrons or holes, must surmount in order to flow across the junction. The direction of easy flow is that in which the carriers flow from the semiconductor to the metal. An applied voltage which produces a current flow in this direction reduces the height of the potential hill and allows the carriers to flow more easily to the metal. When the voltage is applied in the opposite direction the height of the barrier which the carriers must surmount in going from the metal into the semiconductor is unchanged, to a first approximation, and the resistance of the contact remains high. A *p*-type semiconductor is positive, an *n*-type negative, relative to the metal, in the direction of easy flow.

Rectifying contacts can also be made between two semiconductors of opposite conductivity types. The direction of easy flow is again that for which the *p*-type is positive, the *n*-type negative. The rectifying boundary may separate two regions with different conductivity characteristics within the same crystal.

In some contact rectifiers it is necessary to consider the flow of both types of carriers, electrons and holes, even though one type is overwhelmingly in excess under equilibrium conditions. An example is the germanium point contact rectifier such as the 400 type varistor. The germanium used is n-type and the normal concentration of holes is small compared to the concentration of conduction electrons. Nevertheless, a large part of the current in the forward direction consists of holes flowing away from the contact rather than electrons flowing in. The flow increases the concentration of holes in the vicinity of the contact and there is a corresponding increase in the concentration of electrons to compensate for the space charge of the holes. This increase in concentration of carriers increases the conductivity of the germanium. The holes introduced in this way gradually combine with electrons and disappear so that at large distances the current consists largely of electrons. Similar effects occur at *n-p* boundaries in germanium; the current in the forward direction consists in part of holes flowing from the *p*-type region into the *n*-type region and electrons flowing from the *n*-type region into the *p*-type region.

The alteration of concentration of carriers and conductivity by current flow may be used to produce amplification in a number of ways. In the type-A transistor two point contacts are placed in close proximity on the upper face of a small block of *n*-type germanium. A large area low resistance contact on the base is the third element of the triode. Each point, when connected separately with the base electrode, has characteristics similar to those of the rectifier. When operated as an amplifier, one point, called the emitter, is biased in the forward direction so that a large part of the current consists of holes flowing away from the contact. The second point, called the collector, is biased in the reverse direction. In the absence of the emitter, the current consists largely of electrons flowing from the collector point to the base electrode. When the two points are in close proximity there is a mutual influence which makes amplification possible. The collector current produces a field which attracts the positively charged holes flowing from the emitter, so that a large part of the emitter current flows to the collector and into the collector circuit.

It has been found that rectifying boundaries between *n*- and *p*-type germanium may be used both as emitters and collectors, so that it is possible to make transistors without point contacts.

The following five papers are concerned with the behaviors of holes and electrons in semiconductors, with particular emphasis upon rectifying junctions and transistors. The first paper "Hole Injection in Germanium" describes new experiments on the behavior of holes and shows how their numbers and velocities may be measured and how they may be used to modulate the conductivity in the "filamentary transistor." The second paper "Some Circuit Aspects of the Transistor" describes the characteristics and equivalent circuits for the transistor. "Theory of Transient Phenomena in the Transport of Holes in an Excess Semiconductor" describes in mathematical terms a number of the processes encountered in the first paper and brings out interesting features of the nature of an advancing wave front of holes. "The Theory of Rectifier Impedances at High Frequencies" analyzes the behavior of metal-semiconductor rectifiers for high frequencies for the case in which the current is carried by one type of carrier only. As mentioned above, in rectifiers formed from *p-n* junctions, currents of both holes and of electrons must be considered. Such rectifiers and related subjects are dealt with in "The Theory of *p-n* Junctions in Semiconductors and *p-n* Junction Transistors."